

Evaluating Adaboost and Bagging Methods for Time Series Forecasting EEG Dataset

N. Geethanjali, G.T.Prasanna kumari, M.Usha Rani

Abstract. Time series forecasting is a paramount range from claiming machine learning that is frequently neglected. It is critical a direct result there are thus large portions prediction issues that include a period part. These issues are dismissed on account of it, this period part will lead to time series issues more troublesome to manage. An fascinating time series classification issue will be foreseeing if an subject's eyes need aid open alternately shut based best for their brain wave information (EEG). We will aggravate examination for Adaboost and Bagging methodologies on EEG dataset.

Keywords: Adaboost, Bagging, EEG.

1. INTRODUCTION

Time does assume a part to typical machine learning datasets. Predictions would produced to new. The real prediction might not a chance to be known until a future date. What's to come is, no doubt, predicted. Anyway all former perceptions need aid practically constantly dealt with just as. Maybe for a portion exact. Minor transient progress should succeed the clue from claiming "concept drift" for example, just utilizing the most recent quite a while. For perceptions instead of the greater part information accessible. A period arrangement dataset may be separate. Time arrangement includes a unequivocal request reliance between observations: a period measurement and extra extent. May be both a demand and a structure that gives an wellspring for extra majority of the data. A period arrangement may be an arrangement about perceptions taken consecutively in time.

We have different objectives relying upon if we would intrigued by focusing on a dataset or performing final predictions. Focusing on a dataset, known as time series analysis, assistance with settle on finer predictions, may be not required and could bring about an expansive specialized foul financing in time furthermore adroitness not straightforwardly adjusted for the fancied outcome, which is determining what's to come. Clinched alongside spellbinding modeling, alternately time arrangement analysis, a period arrangement will be modeled to figure out its parts as far as occasional patterns, trends, connection to outer factors. For contrast, time arrangement determining utilization the data done a period arrangement (perhaps with extra information) with figure future qualities for that arrangement.

At utilizing established statistics, the essential concern may be the investigation from claiming time arrangement.

Revised Manuscript Received on February 11, 2019.

Dr. N. Geethanjali, Professor, Department of Computer Science, S.K.University, Ananthapur, India (geethanjali.sku@gmail.com)

G.T.Prasanna kumari, Research Scholar, Department of Computer Science and Engineering, S V Engineering College for Women, Tirupati, India (tabithaprasanna@gmail.com)

Dr. M.Usha Rani, Professor, Department of Computer Science, SPMVV, Tirupathi, India (musha_rohan@yahoo.com)

Time arrangement dissection includes creating models that best catch alternately depict a watched time arrangement. In place should comprehend those reasons. This field of focusses on "why" behind a period arrangement dataset which frequently includes performing presumptions something like the structure of the information and disintegrating the time arrangement under constitution parts. Those nature of a spellbinding model may be resolved toward how great it portrays all accessible information and the understanding it gives should exceptional advise the issue web-domain. Those essential objective for time arrangement examination is on create scientific models that gatherings give possible portrayals from example data.

2. TIME SERIES FORECASTING

Making predictions over future data may be known as extrapolation in the traditional statistical handling of about time arrangement information. More up to date fields concentrate on those subject sentence and allude to it similarly as time arrangement determining. Determining includes taking models fit looking into authentic information also utilizing them will foresee future perceptions. Descriptive models might acquire for what's to come (i.e. should smooth birch alternately uproot noise), they best try to best depict those information. A paramount refinement clinched alongside determining may be that what's to come will be totally inaccessibility furthermore must best make assessed starting with what need recently happened. The reason for time arrangement dissection will be by twofold: to see all the or model the stochastic instrument that provides for ascent on a watched arrangement and should anticipate alternately conjecture what's to come qualities of a arrangement in view of those history from claiming that arrangement. Those ability of a time arrangement determining model may be dictated by its execution in foreseeing what's to come. This may be frequently toward those out from claiming having the ability to illustrate the reason a particular prediction might have been made, certainty intervals furthermore actually finer seeing the reasons behind the issue.

3. CONSTITUENT PARTS OF TIME SERIES

Time series analysis has a number of subparts for understanding and analyzing the dataset.

Four constituent parts of a time series:

1. **Level.** If it were a straight line, the baseline value for the series.
2. **Trend.** The up's and down's behaviour of the series over time.



3. **Seasonality.** Repeatative patterns(cycles) of behaviour over time.
4. **Noise.** The variation in the observations that cannot be explained by the model.

Level and noise are mandatory and trend and seasonality are optional for all time series. But, important features of time series are trends and seasonal variations. These constituent parts are combined together to form time series.

4. ELECTROENCEPHALOGRAPHY (EEG)

Assessing machine learning models on time arrangement determining issues is testing. It will be not difficult should aggravate a little lapse in the encircling of a issue alternately in the assessment about models that provide for amazing comes about anyhow bring about an invalid finding. An intriguing time arrangement issue may be foreseeing if an subject’s eyes are open alternately shut based main on their cerebrum wave information (EEG).

We will find those issue of foreseeing if eyes need aid open or shut in light of mind waves furthermore an as a relatable point methodological trap the point when assessing time arrangement determining models. We are setting off will make a closer take a gander at an issue that includes foreseeing if the subjects eyes need aid open or shut in view of cerebrum wave information. Those issue might have been portrayed and information gathered toward Oliver Rosler and David Suendermann for their paper titled “A First Step towards Eye State Prediction Using EEG“. Specifically, an electroencephalography (EEG) recording might have been constructed of a solitary man to 117 seconds (just under two minutes) same time those subject opened and shut their eyes, which might have been recorded by means of a feature polaroid. Those open/closed state might have been after that recorded against each the long run venture in the EEG follow manually. Those EEG might have been recorded utilizing an emotiv EEG Neuroheadset, bringing about 14 follow. Those yield variable is binary, significance that this may be a two-class arrangement issue. What added up to 14,980 perceptions (rows) were made through those 117 seconds, significance that there were over 128 perceptions for every second. Those corpus comprises about fourteen thousand instances with fifteen features every (14 features to the values of the electrodes and the eye state). The instances would saved in the corpus clinched alongside ordered request will have the ability will dissect fleeting dependencies. 8,255(55 %) instances of the corpus relate of the eye state open and 6,722(44%) instances of the eye state shut.

The analyze might have been conveyed out in a tranquil space. Throughout those experiment, those proband might have been being videotaped. To forestall artifacts, the proband might have been not mindful of the correct start period of the estimation. Instead, he might have been totally straightforward with sitting relaxed state, gaze straight of the camera, and transform those eye state in freedom of thought. Best extra demand might have been that, gathered in those whole session, the span about both eye states ought further bolstering make over the same furthermore that the distinct intervals ought further bolstering change incredibly long (from eye squinting will more drawn out stretches) to Figure 1. The span of the estimation might have been 117 seconds. The inspecting rate of the EEG headset A/D converter might

be four times the span rate of the feature polaroid. Those eye state might have been manually annotated and examining the feature recordings adjusted for those EEG information. Both open or incompletely open eyes were sorted as open state; best totally shut eyes were sorted similarly as shut.

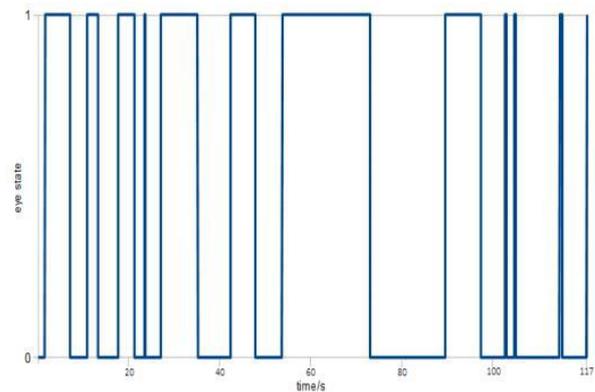


Figure 1. Eye states(for duration of two minutes). 0 represents open and 1 represents closed eye state.

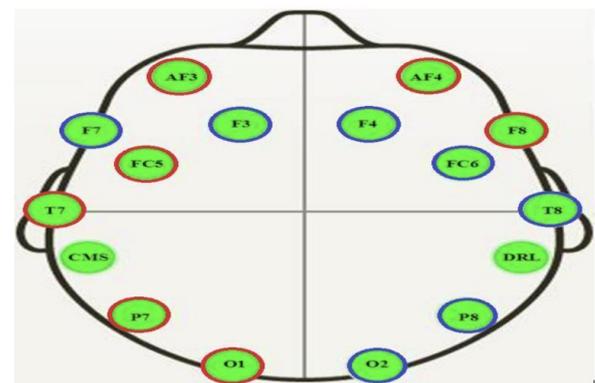


Figure 2. Sensor position; Blue represents to a maximum increase and red represents to a minimum decrease with opening eyes.

Those yield variable may be binary, significance that this may be an two-class arrangement issue. What added up to 14,980 perceptions (rows) were committed again the 117 seconds, significance that there were around 128 perceptions for every second. There were also a portion EEG perceptions that need an significantly bigger over anticipated plentifulness. These need aid liable outliers and a chance to be identifier and evacuated utilizing a straightforward factual system for example, such that uprooting rows that have a perception 3-to-4 standard deviations. Those simplest encircling of the issue will be should anticipate the eye-state (open/closed) provided for the EEG. All the more propelled framings of the issue might try to model the multivariate time arrangement of every EEG follow in place on anticipate the present eye state. The corpus comprises for 14,977 instances for 15 qualities. The instances saved in the corpus over ordered request to have the capacity to dissect fleeting dependencies. 8,255 (55%) instances of the corpus relate of the eye open furthermore 6,722 (44%) instances of the eye shut state. The dataset downloaded free of charge:



• EEG Eye State Data Set

RESULTS & DISCUSSIONS

Relaton: eeg	Nbr	AF3	F7	F3	FC5	T7	P7	O1	O2	P8	T8	FC6	F4	F8	AF4	eyedetction
	Numeric															
1	4329.23	4009.23	4289.23	4148.21	4350.26	4586.15	4096.92	4641.03	4222.05	4238.46	4211.28	4280.51	4635.9	4393.85	4393.85	0
2	4324.62	4004.62	4293.85	4148.72	4342.05	4586.67	4097.44	4638.97	4210.77	4226.67	4207.69	4279.49	4632.82	4384.10	4384.10	0
3	4327.69	4006.67	4295.38	4156.41	4336.92	4583.9	4096.92	4630.26	4207.69	4222.05	4206.67	4282.05	4628.71	4389.23	4389.23	0
4	4328.72	4011.79	4296.41	4155.9	4343.59	4582.56	4097.44	4630.77	4217.44	4235.38	4210.77	4287.69	4632.31	4396.41	4396.41	0
5	4326.15	4011.79	4292.31	4151.28	4347.69	4586.67	4095.9	4627.69	4210.77	4244.1	4212.82	4288.21	4632.82	4398.46	4398.46	0
6	4321.03	4004.62	4289.41	4153.23	4345.64	4587.18	4093.33	4616.92	4202.56	4232.92	4209.74	4281.03	4628.21	4389.74	4389.74	0
7	4318.49	4001.03	4280.51	4151.79	4343.59	4584.62	4089.74	4615.41	4212.31	4236.67	4201.03	4269.74	4625.1	4378.46	4378.46	0
8	4325.64	4006.67	4278.46	4143.08	4344.1	4583.08	4087.18	4614.87	4205.64	4230.26	4195.9	4266.67	4622.05	4380.51	4380.51	0
9	4326.15	4010.77	4276.41	4139.49	4345.13	4584.1	4091.28	4608.21	4187.69	4229.74	4202.05	4273.85	4627.18	4389.74	4389.74	0
10	4326.15	4011.28	4278.92	4142.05	4344.1	4582.56	4092.82	4608.72	4194.36	4228.72	4212.82	4277.69	4627.49	4393.23	4393.23	0
11	4326.15	4010.77	4272.82	4143.08	4342.56	4579.49	4087.69	4615.9	4206.15	4228.72	4210.77	4272.82	4631.79	4382.56	4382.56	0
12	4316.92	4002.56	4259.49	4135.9	4339.49	4581.54	4086.15	4615.38	4195.9	4223.59	4197.44	4262.05	4613.33	4370.77	4370.77	0
13	4308.21	3993.33	4247.69	4124.62	4335.9	4584.62	4089.74	4610.26	4188.72	4221.54	4190.26	4255.9	4606.15	4369.74	4369.74	0
14	4308.21	3996.46	4251.79	4122.56	4338.46	4584.62	4091.28	4618.92	4205.64	4225.54	4194.36	4263.59	4610.26	4371.28	4371.28	0
15	4312.82	4006.15	4260.51	4127.69	4341.03	4586.15	4089.74	4628.72	4215.9	4225.13	4198.97	4274.36	4611.79	4375.92	4375.92	0
16	4313.85	4003.59	4255.9	4127.69	4337.95	4586.15	4091.28	4627.69	4206.15	4224.62	4199.49	4269.23	4610.26	4378.97	4378.97	0
17	4306.67	3997.95	4247.18	4122.56	4336.92	4584.1	4094.36	4620.51	4202.56	4226.15	4193.85	4254.36	4605.13	4369.74	4369.74	0
18	4303.59	3996.41	4249.21	4122.05	4341.03	4581.54	4096.77	4614.41	4203.59	4226.15	4199.23	4253.85	4605.13	4370.26	4370.26	0
19	4312.31	4000.0	4255.9	4124.1	4342.05	4586.67	4085.64	4610.77	4192.31	4223.08	4192.31	4264.62	4614.87	4382.56	4382.56	0
20	4315.9	4003.59	4257.44	4122.56	4341.54	4588.21	4088.21	4607.18	4188.21	4216.92	4200.0	4270.26	4621.54	4378.97	4378.97	0
21	4314.36	4005.64	4259.49	4123.59	4345.13	4588.21	4092.31	4612.82	4205.67	4222.56	4209.74	4272.31	4624.62	4381.54	4381.54	0
22	4324.62	4012.82	4263.08	4132.31	4343.08	4587.18	4091.79	4617.95	4213.33	4240.51	4217.95	4274.36	4632.31	4400.51	4400.51	0
23	4331.28	4024.1	4266.15	4140.0	4336.41	4585.64	4089.74	4617.95	4204.62	4245.13	4216.41	4274.36	4638.97	4401.54	4401.54	0
24	4326.67	4020.0	4264.1	4137.44	4334.36	4586.15	4088.21	4621.54	4213.33	4238.46	4208.72	4271.28	4634.36	4385.13	4385.13	0
25	4321.54	4007.69	4260.0	4132.31	4340.0	4585.64	4087.18	4620.26	4203.59	4242.56	4208.72	4270.26	4627.69	4378.97	4378.97	0
26	4317.44	4007.69	4257.44	4128.72	4343.59	4585.13	4088.21	4638.97	4211.79	4247.69	4209.23	4269.74	4625.13	4386.15	4386.15	0
27	4311.79	4007.18	4250.77	4125.13	4336.41	4587.18	4087.69	4634.36	4209.23	4232.31	4197.44	4262.05	4615.38	4381.03	4381.03	0
28	4308.72	4000.0	4245.13	4121.54	4329.23	4585.64	4086.67	4617.44	4189.23	4216.41	4186.15	4253.85	4604.1	4362.05	4362.05	0
29	4308.21	4002.05	4242.82	4122.56	4332.82	4582.56	4087.69	4610.77	4192.31	4216.41	4199.74	4256.92	4605.13	4360.0	4360.0	0
30	4311.28	4011.28	4265.13	4130.26	4339.49	4583.59	4089.74	4616.41	4200.0	4220.51	4195.9	4265.64	4611.28	4374.36	4374.36	0
31	4311.28	4008.72	4265.13	4130.77	4337.44	4585.64	4087.69	4614.87	4190.26	4218.46	4195.38	4265.64	4608.72	4374.36	4374.36	0
32	4306.15	4000.51	4259.49	4121.03	4330.26	4584.1	4085.64	4607.18	4185.13	4216.92	4200.0	4263.08	4611.28	4369.23	4369.23	0
33	4307.18	4003.08	4263.08	4121.54	4330.77	4582.56	4086.67	4611.79	4202.05	4224.62	4208.72	4271.28	4622.56	4379.49	4379.49	0

Table 1. EEG ARFF File.

The Table 1 comprises for 14977 instances with 15 qualities. The instances are saved in the corpus in which 8255 (55%) instances of the corpus relate of the eye open furthermore 6722 (44%) instances of the eye shut state. Table 2 indicates the quality ranges of the 14 sensors in the corpus. There is an self-evident distinction on plentifulness about specific sensors the point when analyzing those go about qualities to diverse eye states. Those sensors F7, F3, O2, P8, T8, FC6, and F4, those greatest values for the eye open state are higher over values of the eye shut state same time the base qualities would about the same. On the other hand, to those sensors AF3, FC5, T7, P7, O1, F8, and AF4, those least values to those eye open state would more level over to those eye shut state same time those most extreme values are something like the same. Constantly on sensors need on regular that open eye state goes alongside an higher quality extent over those eye shut state same time the intend sits tight almost the same. Accordingly, additionally those standard deviation increments. Sensors might make part under two assemblies. In the primary group, the greatest increments when eyes open while, in the other group, the least declines in the same. A large portion sensors of the primary gathering happen with make spotted on the right side, same time the vast majority of the second assembly need aid on the left side of the brain, similarly as shown previously, Figure 2.

Eye State	closed			open		
	min	mean	max	min	mean	max
AF3	4198	4305	4445	1030	4297	4504
F7	3905	4005	4138	3924	4013	7804
F3	4212	4265	4367	4197	4263	5762
FC5	4058	4121	4214	2453	4123	4250
T7	4309	4341	4435	2089	4341	4463
P7	4574	4618	4708	2768	4620	4756
O1	4026	4073	4167	3581	4071	4178
O2	4567	4616	4695	4567	4615	7264
P8	4147	4202	4287	4152	4200	4586
T8	4174	4233	4323	4152	4229	6674
FC6	4130	4204	4319	4100	4200	5170
F4	4225	4281	4368	4201	4277	7002
F8	4510	4610	4811	86	4601	4833
AF4	4246	4367	4552	1366	4356	4573

Table 2. Ranges And Means Of The Sensor Values For The Eye States.

5. COMPARISON OF ADABOOST AND BAGGING ALGORITHM ON EEG DATASET

For implementing AdaBoost and Bagging approaches, we use the Weka Toolkit. We used two-fold cross-validation with their default parameter settings. Results of this experiment are shown in Figure 5. Surprisingly, Adaboost classifier with a high classification performance(89%) produced compared with Bagging. For AdaBoost an error rate of 0.1 and time taken to build model is 104 seconds.

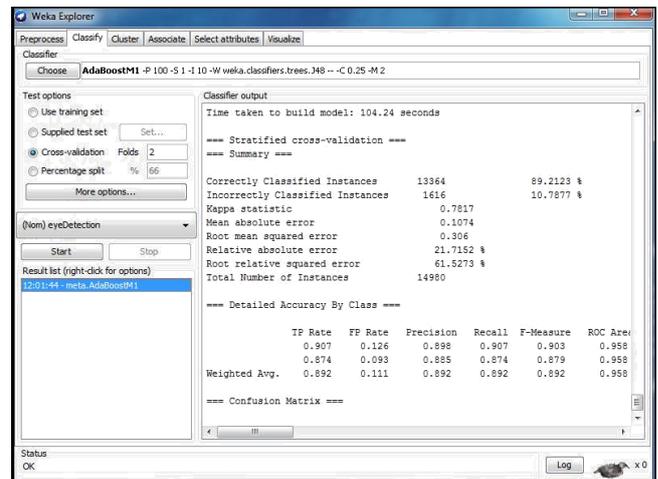


Figure 3. Eye State Classification using Adaboost Algorithm.

Bagging got an error rate of 0.2 and processing time to build model is 45 seconds.

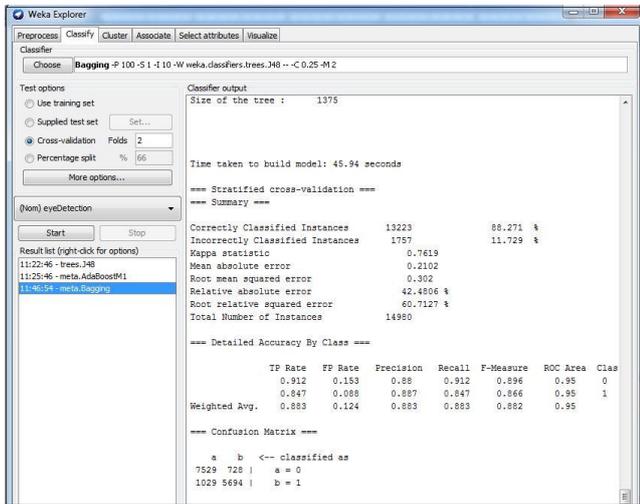


Figure 4. Eye State Classification using Bagging Algorithm.

	ERROR(%)	CLASSIFIED PERFORMANCE (%)	TIME TAKEN TO BUILD THE MODEL(S)
ADABOOST	0.11	89	104.24
BAGGING	0.21	88	45.94

Figure 5. Analysis of Adaboost and Bagging Approaches.

After running two different Adaboost and Bagging approaches, we achieved a classification error rate of 0.1% which is better than error rate of Bagging. We have demonstrated comparative study of AdaBoost and Bagging classification algorithms with the help of WEKA. Based on the accuracy and classification performance point of view, AdaBoost works better than Bagging algorithm. This shows that it is possible to predict eye state using EEG sensor input with an high accuracy.

6. CONCLUSION

In this paper, we presented an dissection of classification performance and error rate of EEG dataset using AdaBoost and Bagging approaches. We showed similar contemplate about AdaBoost and and Bagging algorithms with those assistance from claiming WEKA. The test outcomes indicate that, for the exactness furthermore transforming chance perspective from claiming view, AdaBoost meets expectations superior to Bagging algorithm. AdaBoost approach serves the time arrangement order issue done foreseeing if a subject’s eyes would open or shut built just looking into their brain wave information (EEG).

REFERENCES

1. Agriculture soil classification and fertilizer recommendation using Adaboost and Bagging approaches: G.T.Prasanna Kumari; SSRN, ELSEVIER, January 2018.
2. Analysis of agriculture data using data mining techniques: application of big data: Jharna Majumdar, Sneha Naraseeyappa and Shilpa Ankalaki; Journal of Big Data, Majumdaret al. *J Big Data (2017) 4:20* DOI 10.1186/s40537-017-0077-4.
3. An Overview of Data Mining Techniques Applied to Agricultural Soil Data: BalajiAglave, Razaullah Khan, Sanjay Sirsat; International Journal of Agriculture Innovations and Research Volume 3, Issue 2, ISSN (Online) 2319-1473, 11/09/2014.
4. A Study of AdaBoost and Bagging Approaches on Student Dataset: G. T. Prasanna Kumari, Dr. M. Usha Rani; International Research Journal of Advanced Engineering and

Science, Volume 2, Issue 2, pp. 375-380, 2017, ISSN (Online): 2455-9024.

5. An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization: THOMAS G. DIETTERICH; Machine Learning, 40, 139–157, Kluwer Academic Publishers. Manufactured in The Netherlands.
6. Multiple classifier system for EEG signal classification with application to brain–computer interfaces: Amir Ahangi, Mehdi Karamnejad, Nima Mohammadi, Reza Ebrahimpour, Nasoor Bagheri; Neural Comput & Applic (2013) 23:1319–1327 DOI 10.1007/s00521-012-1074-3.
7. A Data Mining Based On Ensemble Classifier Classification Approach for Edible Mushroom Identification: Muhammad Husaini; International Research Journal of Engineering and Technology (IRJET), Volume: 05 Issue: 07 | July 2018.
8. Covariate Shift Estimation based Adaptive Ensemble Learning for Handling Non-Stationarity in Motor Imagery related EEG-based Brain-Computer Interface: Haider Raza and Shang-Ming Zhou; Preprint submitted to Elsevier, May 4, 2018.
9. An Ensemble of Classifiers based Approach for Prediction of Alzheimer’s Disease using fMRI Images based on Fusion of Volumetric, Textural and Hemodynamic Features: Fatima MALIK, Saima FARHAN, Muhammad Abuzar FAHIEM; Advances in Electrical and Computer Engineering Volume 18, Number 1, 2018.